REMARKS/ARGUMENTS

I. Introduction:

Claims 1-38 are currently pending.

II. Claim Rejections under 35 U.S.C. 103:

Claims 1, 6-8, 10, 12, 14-15, 17, 21-29, 31, 34, and 36 stand rejected under 35 U.S.C. 103(a) as being unpatentable over U.S. Patent No. 6,163,796 (Yokomizo) in view of U.S. Patent No. 6,351,775 (Yu) and further in view of U.S. Patent No. 6,801,949 (Bruck et al.).

Yokomizo is directed to a network system having a plurality of multimedia servers for processing different types of data. As illustrated in Fig. 2, the system includes a center server in communication with a plurality of function servers, configured to perform different functions. Applicants respectfully submit that Yokomizo does not disclose selecting a real server for connection with a client, sending a redirect message to the client specifying the selected real server, or receiving a new connection request from the client for connection with the selected real server, as set forth in the claims.

The client of Yokomizo does not connect directly to the function server, thus there is no selection of a real server for connection with the client. When a client requests a service, the client issues a process script to the center server, which in turn accesses one of the function servers. The center server processes the request directly with the function server and then returns a final result to the client. The fact that the client does not have to connect with the function server is an important feature of the Yokomizo system. Since the client only needs to wait for the final result from the center server, the client can execute another operation during this interval to improve throughput. Furthermore, the client does not need to know access methods to the function servers or even their existence. Therefore, there is no

reason to send a redirect message to the client specifying a selected real server or receive a new connection request from the client for connection with a real server.

Furthermore, Yokomizo teaches away from connecting a client to a selected real server for the duration of a transaction. As discussed above, Yokomizo does not want to connect a client directly to one of the function servers, instead the center server connects with one of the function servers. It is an object of Yokomizo to provide a network system that can utilize many function servers to process different types of data. Importantly, the center server is operable to connect with any number of the function servers to perform various operations specified by the process script. Thus, Yokomizo teaches away from a persistent connection between the center server and one of the function servers for the duration of a transaction, or script. If the center server had a persistent connection with one of the function servers, it would defeat the purpose of providing one server that can interact with different function servers to provide different processing.

Yu is directed to load balancing across servers and uses a partioning method to group object identifiers into classes. Clients are required to maintain a class-to-server assignment table to map each class into a server selection. The class-to-server assignment tables can change dynamically at any time as the workload varies. An important aspect of the Yu invention is that the system not only balances a load across a cluster of servers, but also optimizes a cache hit ratio in a given server by localizing identical object requests. Routing requests for the same object to a single server node results in a better cache hit probability at the same server node. In order to optimize cache hit, object identifiers are grouped into classes which are identified in the class-to-server assignment table to map each class into a server selection. The class-to-server assignment table assigns each class to a virtual server and routers dynamically map each virtual server to one of the real servers in the cluster.

Applicants respectfully submit that Yu does not disclose binding a primary virtual server to a set of real servers, as set forth in the claims. Applicants' invention provides binding between a primary virtual server to a set of URLs, each URL having an associated

real server. In contrast, Yu groups map object identifiers into classes and assigns each class to a virtual server. A router then dynamically maps each virtual server to one of the actual servers in a cluster. As noted at col. 11, lines 61-62, the number of virtual servers is greater than the actual number of servers in the server cluster. Since each client in the load balancing system of Yu maintains a class-to-server assignment table, there is no need to send a request for connection to a primary virtual server. The load balancing is performed by updating the assignment table (col. 6, lines 30-33). The specific server to which an object request is sent is determined from the assignment table. Thus, Yu teaches away from sending a request to a primary virtual server. Similarly, since the server information is contained within the assignment table, there is no need to send a redirect message to the client specifying a selected real server.

Furthermore, Yu monitors the load of each server and dynamically updates the class-to-server assignment to improve load balancing. These updates can be provided at any time, including while a client and server are conducting a session or transaction. (See, for example, discussion of reassignment routine at col. 9, lines 63-65, and col. 10, lines 4-42, and col. 12, lines 36-48 "if a server receives a request from a requester that is no longer assigned to that server, the server will inform the requester of the server to which future requests should be issued."). The system of Yu can update assignment tables every minute to assign the requester to a less loaded server (col. 12, lines 49-63). Thus, a requested node can be connected to a number of different servers during a transaction or session.

Bruck et al. disclose a distributed server cluster with graphical user interface. The system includes a list of virtual IP addresses and corresponding node assignments for those addresses. A sticky virtual IP address assignment is used so that the virtual IP address is forced to an assignment to a particular node, so that all traffic for that virtual IP address must be directed to that node. The Bruck et al. patent simply discloses a persistent connection for an IP address and a specific node and does not address a persistent connection for a specific transaction.

Applicants' invention is particularly advantageous in that it provides a persistent (or "sticky") connection between a user and a server. The sticky connection allows a controller or load balancer to direct each client connection in a session to the same server so that all requests from a given client are redirected to the same server and the client remains attached to a single server for the duration of the session between the client and the server. Since HTTP, for example, does not carry state information for applications such as shopping baskets, financial transactions, or interactive games, it is important for the user to be mapped to the same server for each request until a transaction or session is complete.

Accordingly, claim 1 is submitted as patentable over the prior art of record. Claims 2-12 and claims 24-31, depending either directly or indirectly from claim 1, are submitted as patentable for the same reasons as claim 1.

Claims 7 and 12 are further submitted as patentable over Yu which does not disclose sending an HTTP redirect message to the client specifying the selected real server, or receiving a request at a local director, as set forth in claims 7 and 12, respectively. As discussed above, each requester node of Yu maintains a server assignment table to map each class into a server selection. There is, therefore, no need to send a request to a local director to select a server or send a redirect message to the client to inform the client of the selected server.

Independent claims 14, 21, 22, and 23 are submitted as patentable for the reasons discussed above with respect to claim 1.

Claims 15-17 and 34-36, depending either directly or indirectly from claims 14 and 21, respectively, are submitted as patentable for the same reasons as their respective independent claims.

Claims 2-5, 9, 13, 16, 18-20, 32, 35, 37-38 stand rejected under 35 U.S.C. 103(a) as being unpatentable over Yokomizo in view of Yu and further in view of Bruck et al., and further in view of U.S. Patent No. 6,609,213 (Nguyen et al.). The Nguyen et al. patent was

filed on August 10, 2000. The present patent application was filed on June 30, 2000. Therefore, the Nguyen et al. patent is not prior art with regard to the present application. Applicants request that all rejections based on the Nguyen et al. reference be withdrawn. These claims are also submitted as patentable over Yokomizo, Yu, and Bruck et al., for at least the reasons discussed above with respect to claim 1.

The other references cited, including U.S. Patent No. 6,597,956 (Aziz et al.) and U.S. Patent Publication No. 2001/0052024 (Devarakonda et al.), do not remedy the deficiencies of the primary references.

III. Conclusion:

For the foregoing reasons, Applicants believe that all of the pending claims are in condition for allowance and should be passed to issue. If the Examiner feels that a telephone conference would in any way expedite prosecution of the application, please do not hesitate to call the undersigned at (408) 399-5608.

Respectfully submitted,

Cindy S. Kaplan
Reg. No. 40,043

P.O. Box 2448
Saratoga, CA 95070
Tel: 408-446-8690
Fax: 408-446-8691